IBM PROJECT

RESEARCH AGENT

Presented By:
Saranga Pradeep
Ace Engineering Colllege
Cse (Data Science)

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result (Output Image)
- Conclusion
- Future Scope
- References

PROBLEM STATEMENT

Researchers often spend a lot of time searching for papers, summarizing content, managing citations, and drafting reports. These tasks are repetitive and slow down the research process. There is a need for an AI system that can help with these activities. A Research Agent can understand questions, find relevant papers, create summaries, organize references, and even help write research sections. This makes research faster, more accurate, and more efficient for both academic and industrial use.

PROPOSED SOLUTION

- **The proposed system aims to solve the challenges researchers face in managing academic tasks like literature search, paper summarization, citation organization, and research drafting. To address this, we have developed an AI-based Research Agent using IBM Cloud Lite services and IBM Granite models.**
- **Research Query Understanding::**
  - The agent accepts natural language input from users (e.g., research questions or keywords). Using IBM Granite LLMs, it interprets the query context and prepares it for search and analysis.
- **Document Search & Knowledge Retrieval::**
  - It connects to open-access research repositories such as Semantic Scholar, arXiv, and CrossRef APIs to retrieve relevant academic papers or metadata. Retrieved results are filtered and ranked for relevance.
- **Summarization & Content Analysis::**
  - The Research Agent uses IBM Watsonx and Granite models to summarize uploaded or fetched research papers. It can generate abstract-style summaries, key points, or section-wise breakdowns (e.g., Introduction, Methodology, Conclusion).
- **Deployment:**
  - The solution is hosted using IBM Cloud services. Backend is developed in Python and deployed via IBM Code Engine or IBM Cloud Functions for serverless scalability. Files and documents are stored using IBM Cloud Object Storage.
- **Evaluation:**
  - The system's responses are evaluated based on relevance, summarization quality, and citation accuracy. User feedback is used to improve prompt engineering and fine-tune the response logic.

SYSTEM APPROACH

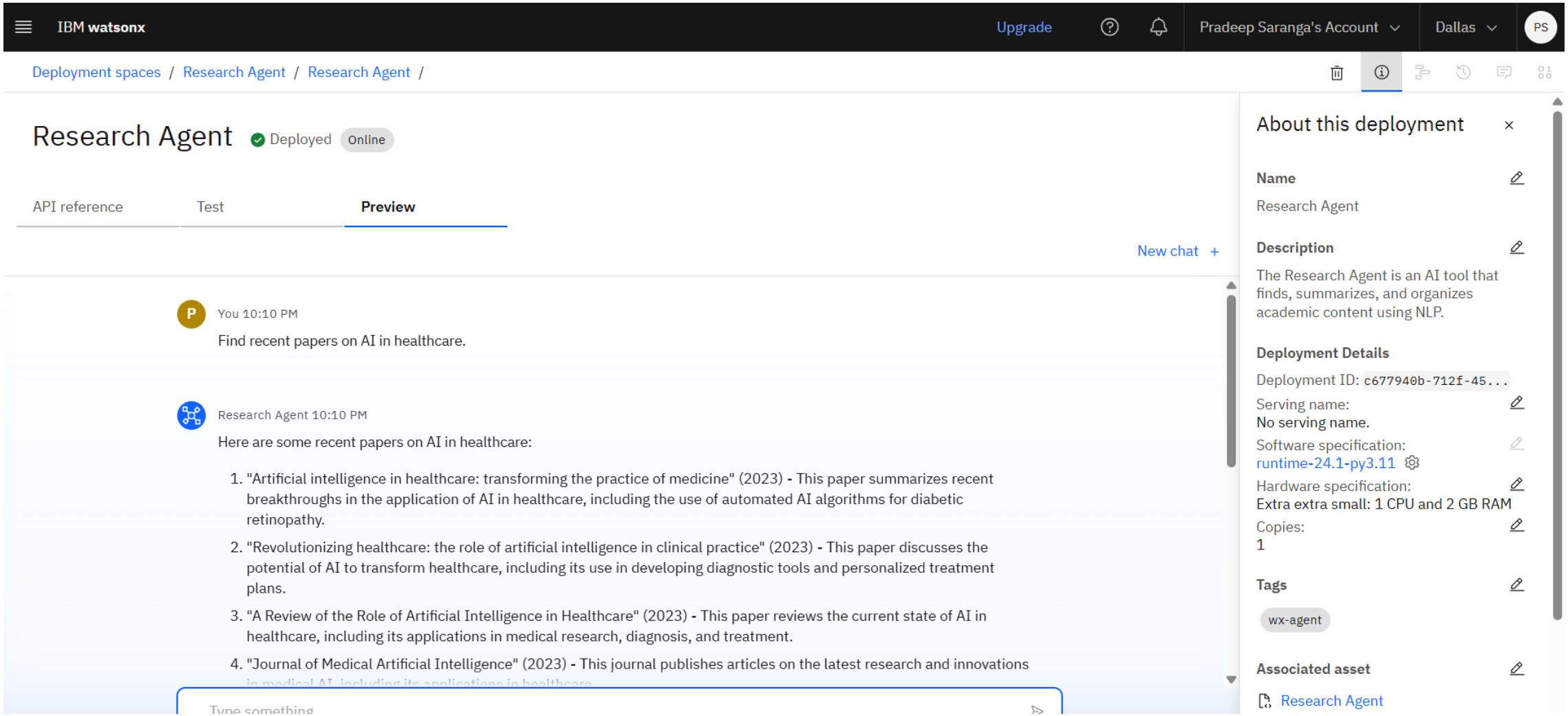
The "System Approach" section outlines the overall strategy and methodology used to design, develop, and deploy the Research Agent system on IBM Cloud. This includes the infrastructure, tools, frameworks, and AI models integrated to achieve the desired functionality.

- **System requirements**
  - **IBM Cloud Lite Account**
  - **IBM Watsonx.ai and IBM Granite Model Access**
  - **IBM Cloud Object Storage** (for file and document storage)
- **Library required to build the model**
  - **IBM Watsonx & Granite** – for NLP tasks like summarization, question answering, and hypothesis generation
  - **Markdown/HTML parsers** – to render formatted summaries or citation outputs on frontend

ALGORITHM & DEPLOYMENT

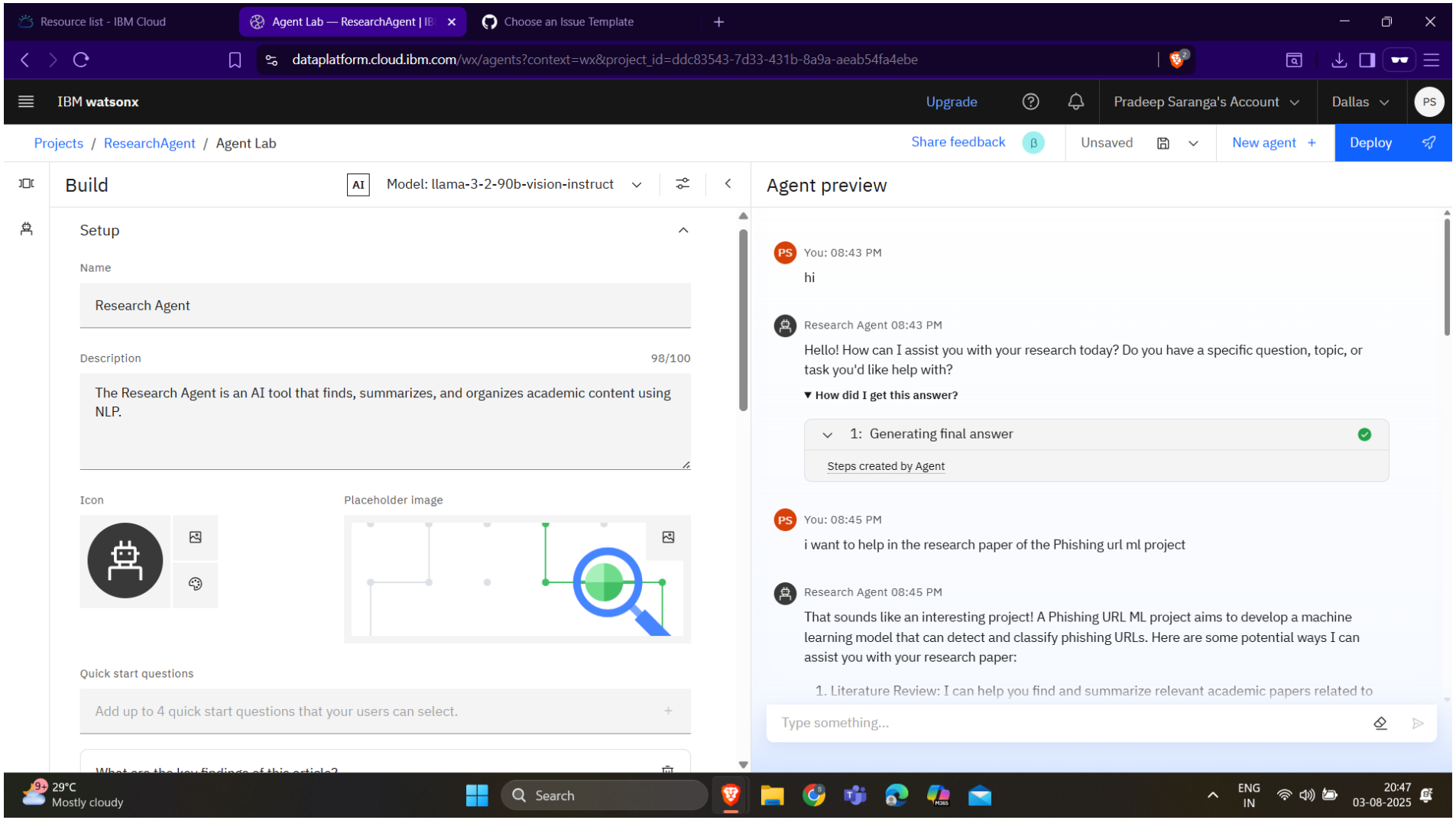
- **Algorithm Selection:**
  - The system uses IBM's **Granite large language models** (LLMs) through Watsonx.ai for natural language processing tasks such as summarization, question answering, citation generation, and hypothesis suggestion. These models were selected for their accuracy in understanding academic language and generating coherent, structured outputs. Unlike traditional ML models, Granite LLMs are pre-trained on vast text corpora and perform well with zero- or few-shot prompting.
- **Data Input:**
  - The inputs to the system include:
  - User queries (natural language research questions)
  - Uploaded research papers (PDF documents converted to text)
- **Training Process:**
  - Extracted text from papers is chunked and embedded using Granite or other embedding models.
  - Vector similarity search is performed to retrieve relevant content based on user queries.
  - Retrieved content is passed as context to prompt templates in Granite for summarization, Q&A, or citation generation.
- **Prediction Process:**
  - Deployed via **IBM Code Engine** or **IBM Cloud Functions** for scalability. File storage is handled using **IBM Cloud Object Storage**
  - Watsonx.ai is used to call Granite LLMs via prompt chaining or REST APIs. The system can be accessed through a simple web interface or chatbot UI.

RESULT



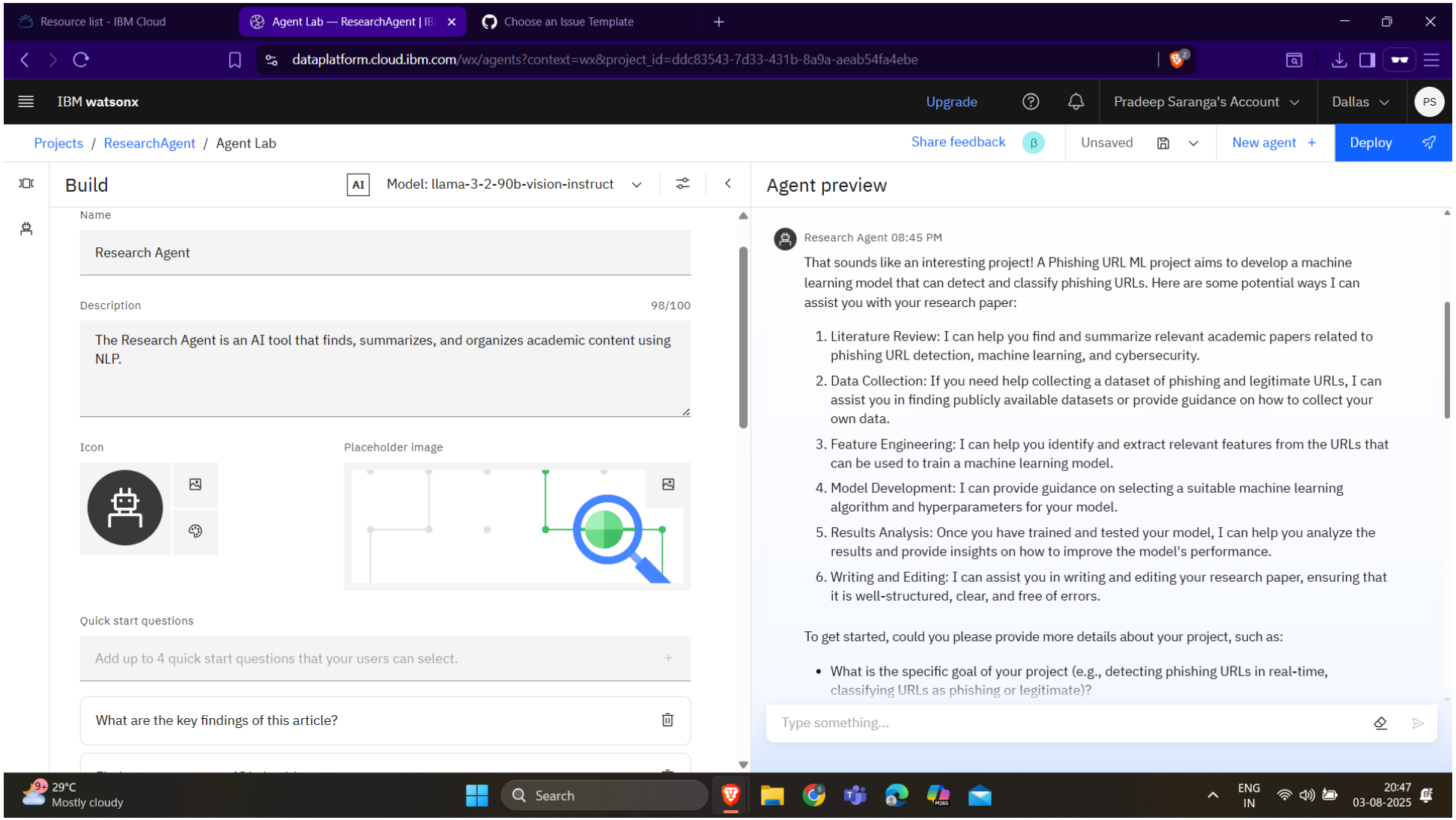
edunet foundation

RESULT



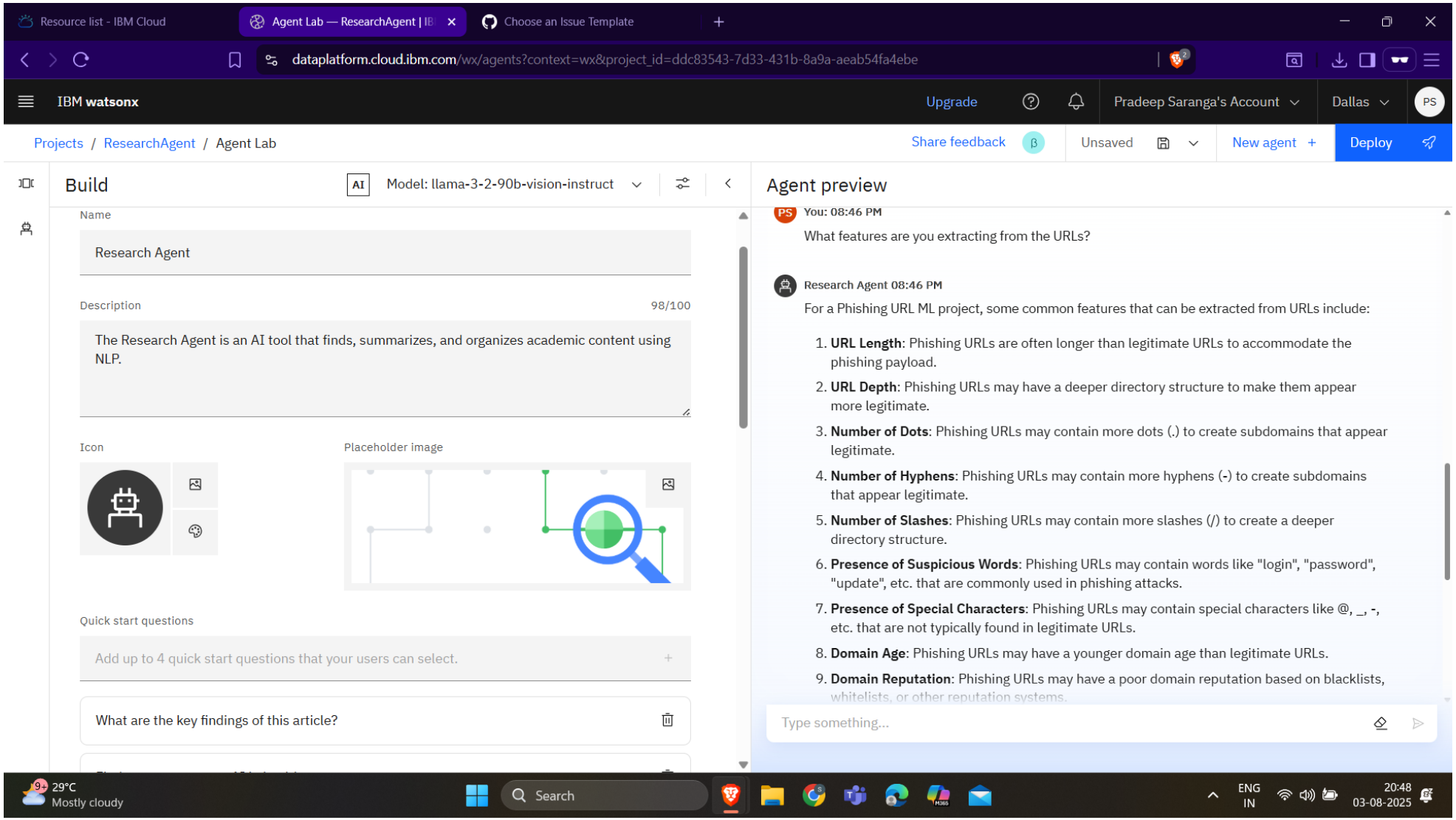
edunet foundation

RESULT



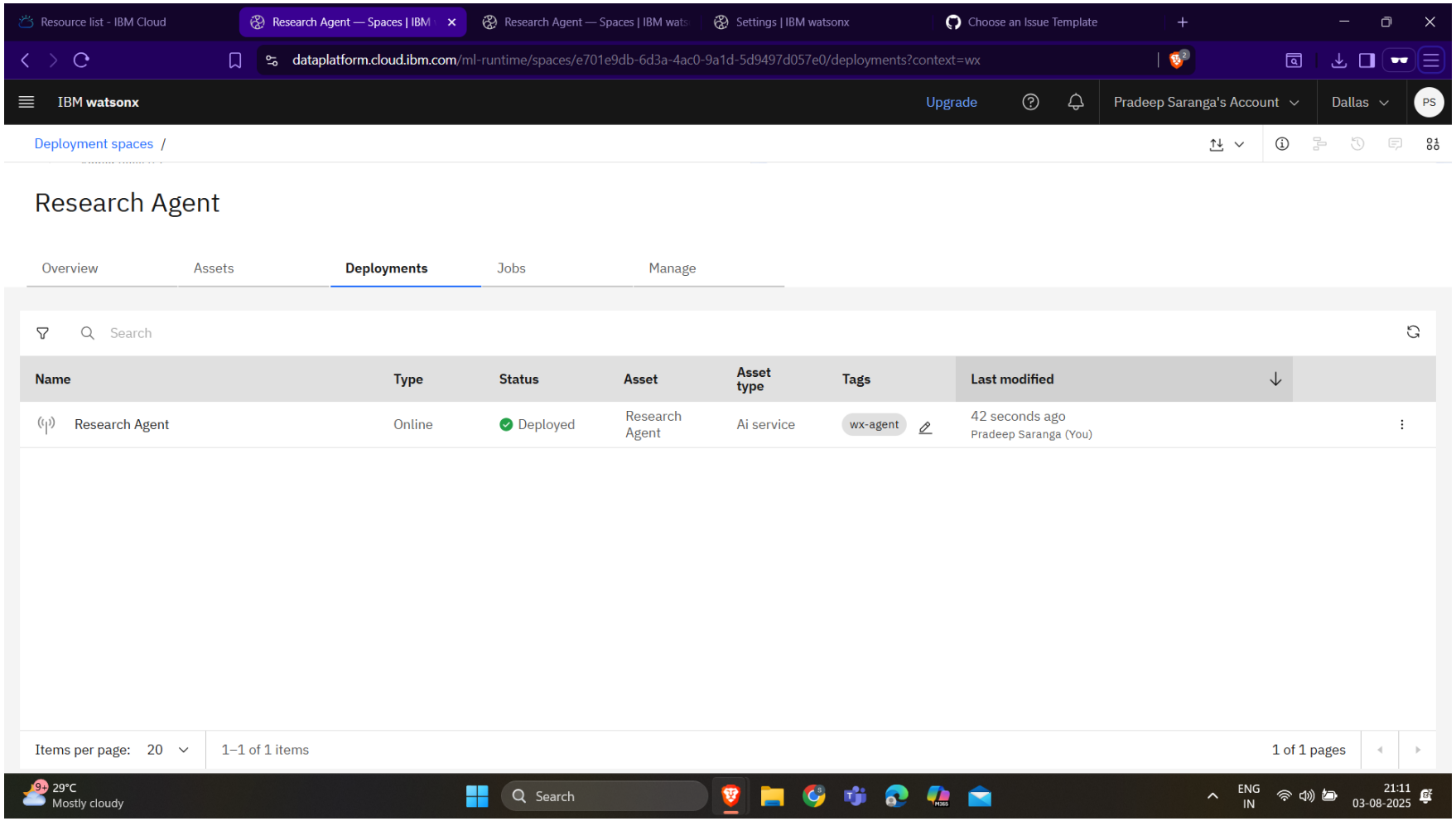
edunet foundation

RESULT



edunet foundation

RESULT



edunet foundation

CONCLUSION

- The Research Agent helps make academic work easier by finding papers, summarizing them, managing citations, and drafting research content. It uses IBM Cloud and Granite models to provide fast and accurate support. The system saves time and improves the quality of research. Some challenges faced were in getting accurate results and handling PDF files, but they were solved using proper tools and prompts. Overall, the Research Agent is a useful tool for students and researchers, and it can be improved further by adding more features in the future.

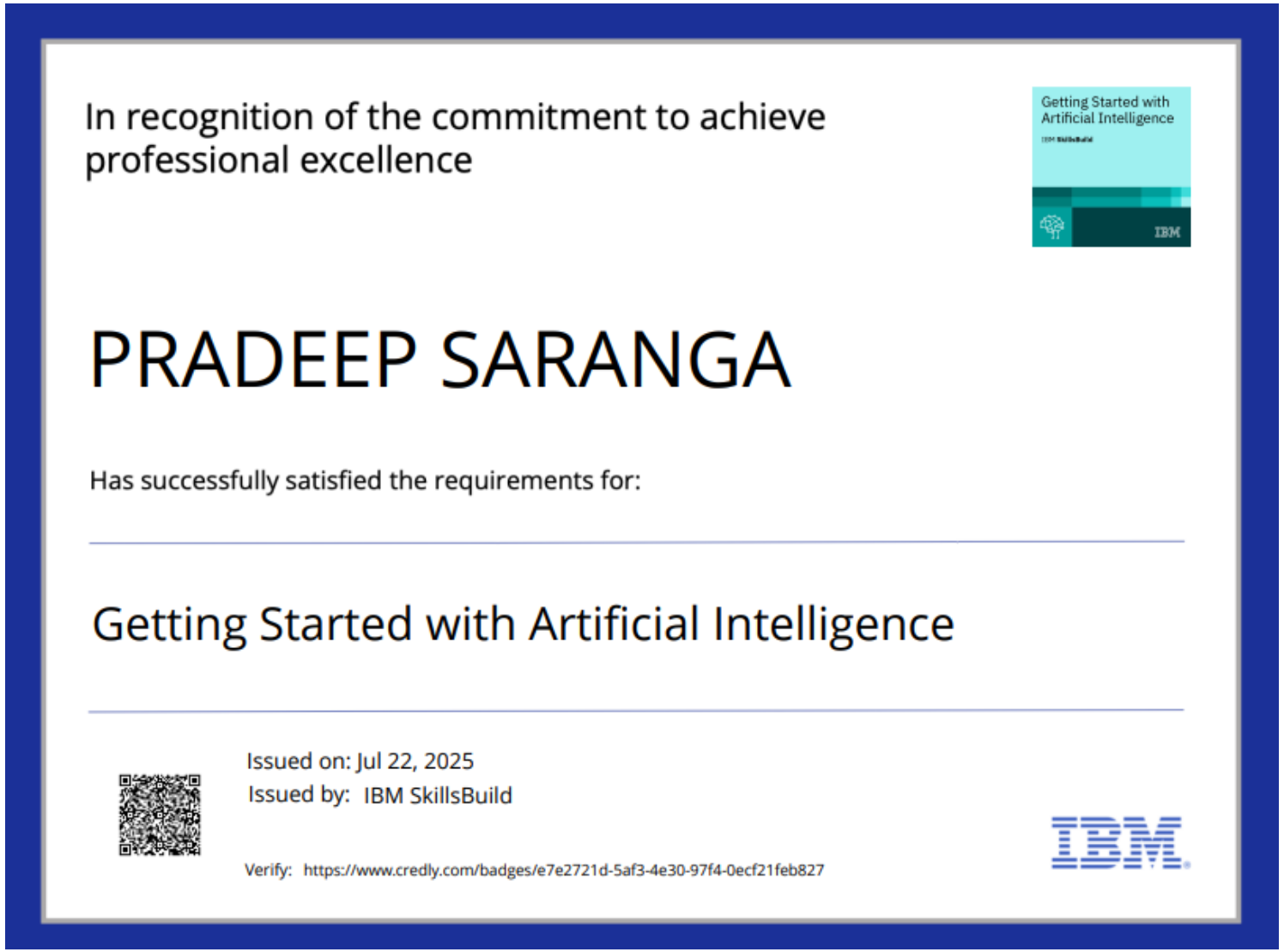
FUTURE SCOPE

- Enable **plagiarism detection** and originality checking for research content.
- Recommend suitable **journals or conferences** for paper submission.
- Introduce **voice input** and **voice response** for hands-free interaction.
- Add **real-time chat interface** for conversational research assistance.
- Allow **collaborative project spaces** for teams working on the same research.
- Include **research deadline reminders** and progress tracking features.
- Expand citation tools to support **LaTeX, BibTeX**, and **Zotero integration**.
- Build a **mobile-friendly version** for access on the go.

REFERENCES

- IBM Watsonx.ai Documentation – *https://www.ibm.com/cloud/watsonx*
- IBM Granite Foundation Models – *https://www.ibm.com/blogs/research/2023/12/ibm-granite-models/*
- PyMuPDF (fitz) Library – *https://pymupdf.readthedocs.io/*

IBM CERTIFICATIONS



IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

PRADEEP SARANGA

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

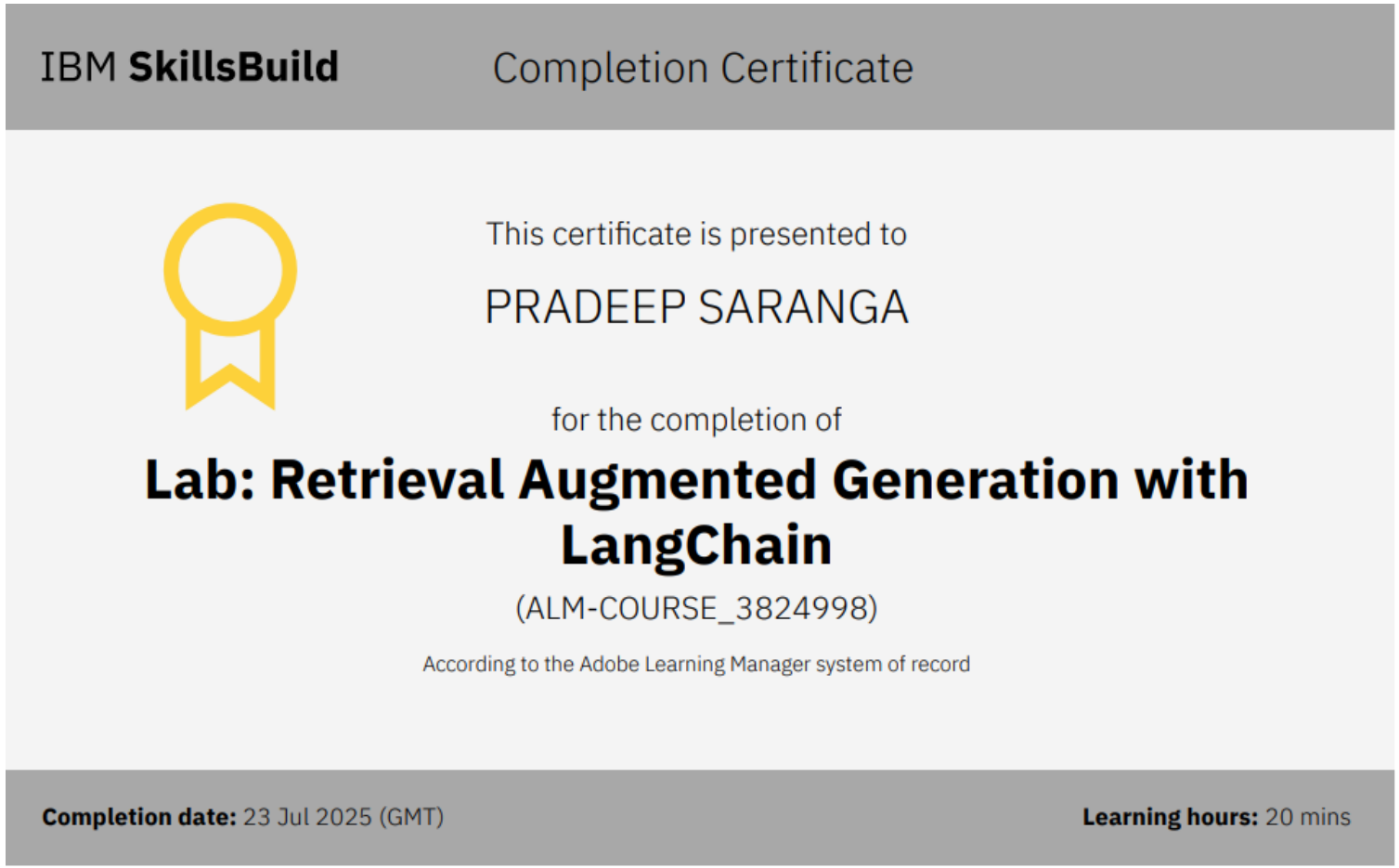
Issued on: Jul 22, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/05bc0fdf-6ce2-41dd-846a-d9cf1e755e64

IBM

IBM CERTIFICATIONS



IBM **SkillsBuild** Completion Certificate

This certificate is presented to

PRADEEP SARANGA

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 23 Jul 2025 (GMT) **Learning hours:** 20 mins

https://skills.yourlearning.ibm.com/certificate/ALM-COURSE_3824998

edunet foundation

THANK YOU

edunet foundation